

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Jarl Wikberg
(B) STREET: Trillvaegen 13
(C) CITY: Umea
(E) COUNTRY: Sweden
(F) POSTAL CODE (ZIP): 905 92 Umea

(A) NAME: Vijay Chhajlani
(B) STREET: Stigbergsvaegen
(C) CITY: Uppsala
(E) COUNTRY: Sweden
(F) POSTAL CODE (ZIP): 752 42 Uppsala

(ii) TITLE OF INVENTION: Human Melanocyte stimulating hormone receptor

(iii) NUMBER OF SEQUENCES: 20

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Sterne, Kessler, Goldstein & Fox
(B) STREET: 1100 New York Ave., N.W.
(C) CITY: Washington
(D) STATE: D.C.
(E) COUNTRY: USA
(F) ZIP: 20005

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/387,805
(B) FILING DATE: 21-FEB-95
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PCT/DK93/00273
(B) FILING DATE: 20-AUG-93

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: DK 1046/92
(B) FILING DATE: 21-AUG-92

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: DK 1118/92
(B) FILING DATE: 10-SEP-92

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: DK 0528/93
(B) FILING DATE: 05-MAY-93

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Cimbala, Michele A.
 (B) REGISTRATION NUMBER: 33,851
 (C) REFERENCE/DOCKET NUMBER: 1102.0160000

(ix) TELECOMMUNICATION INFORMATION:
 (A) TELEPHONE: (202) 371-2600
 (B) TELEFAX: (202) 371-2540

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1270 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (cDNA)

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 169..1122

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGAGAGGGTG TGAGGGCAGA TCTGGGGGTG CCCAGATGGA AGGAGGCAGG CATGGGGGAC	60
ACCCAAGGCC CCCTGGCAGC ACCATGAACT AAGCAGGACA CCTGGAGGGG AAGAACTGTG	120
GGGACCTGGA GGCCTCCAAC GACTCCTTCC TGCTTCCTGG ACAGGACT ATG GCT GTG	177
Met Ala Val	
1	
CAG GGA TCC CAG AGA AGA CTT CTG GGC TCC CTC AAC TCC ACC CCC ACA	225
Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser Thr Pro Thr	
5 10 15	
GCC ATC CCC CAG CTG GGG CTG GCT GCC AAC CAG ACA GGA GCC CGG TGC	273
Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly Ala Arg Cys	
20 25 30 35	
CTG GAG GTG TCC ATC TCT GAC GGG CTC TTC CTC AGC CTG GGG CTG GTG	321
Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu Gly Leu Val	
40 45 50	
AGC TTG GTG GAG AAC GCG CTG GTG GTG GCC ACC ATC GCC AAG AAC CGG	369
Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala Lys Asn Arg	
55 60 65	
AAC CTG CAC TCA CCC ATG TAC TGC TTC ATC TGC TGC CTG GCC TTG TCG	417
Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu Ala Leu Ser	
70 75 80	
GAC CTG CTG GTG AGC GGG AGC AAC GTG CTG GAG ACG GCC GTC ATC CTC	465
Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala Val Ile Leu	
85 90 95	
CTG CTG GAG GCC GGT GCA CTG GTG GCC CGG GCT GCG GTG CTG CAG CAG	513
Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val Leu Gln Gln	

100	105	110	115	
CTG GAC AAT GTC ATT GAC GTG ATC ACC TGC AGC TCC ATG CTG TCC AGC				561
Leu Asp Asn Val Ile Asp Val Ile Thr Cys Ser Ser Met Leu Ser Ser	120	125	130	
CTC TGC TTC CTG GGC GCC ATC GCC GTG GAC CGC TAC ATC TCC ATC TTC				609
Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile Ser Ile Phe	135	140	145	
TAC GCA CTG CGC TAC CAC AGC ATC GTG ACC CTG CCG CGG GCG CGG CGA				657
Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg Ala Arg Arg	150	155	160	
CGC GTT GCG GCC ATC TGG GTG GCC AGT GTC GTC TTC AGC ACG CTC TTC				705
Arg Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser Thr Leu Phe	165	170	175	
ATC GCC TAC TAC GAC CAC GTG GCC GTC CTG CTG TGC CTC GTG GTC TTC				753
Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu Val Val Phe	180	185	190	
TTC CTG GCT ATG CTG GTG CTC ATG GCC GTG CTG TAC GTC CAC ATG CTG				801
Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val His Met Leu	200	205	210	
GCC CGG GCC TGC CAG CAC GCC CAG GGC ATC GCC CGG CTC CAC AAG AGG				849
Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu His Lys Arg	215	220	225	
CAG CGC CCG GTC CAC CAG GGC TTT GGC CTT AAA GGC GCT GTC ACC CTC				897
Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala Val Thr Leu	230	235	240	
ACC ATC CTG CTG GGC ATT TTC TTC CTC TGC TGG GGC CCC TTC TTC CTG				945
Thr Ile Leu Leu Gly Ile Phe Phe Leu Cys Trp Gly Pro Phe Phe Leu	245	250	255	
CAT CTC ACA CTC ATC GTC CTC TGC CCC GAG CAC CCC ACG TGC GGC TGC				993
His Leu Thr Leu Ile Val Leu Cys Pro Glu His Pro Thr Cys Gly Cys	260	265	270	
ATC TTC AAG AAC TTC AAC CTC TTT CTC GCC CTC ATC ATC TGC AAT GCC				1041
Ile Phe Lys Asn Phe Asn Leu Phe Leu Ala Leu Ile Ile Cys Asn Ala	280	285	290	
ATC ATC GAC CCC CTC ATC TAC GCC TTC CAC AGC CAG GAG CTC CGC AGG				1089
Ile Ile Asp Pro Leu Ile Tyr Ala Phe His Ser Gln Glu Leu Arg Arg	295	300	305	
ACG CTC AAG GAG GTG CTG ACA TGC TCC TGG TGAGCGCGGT GCACGCGCTT				1139
Thr Leu Lys Glu Val Leu Thr Cys Ser Trp	310	315		
TAAGTGTGCT GGGCAGAGGG AGGTGGTGAT ATTGTGTGGT CTGGTTCCTG TGTGACCCCTG				1199
GGCAGTTCCT TACCTCCCTG GTCCCCGTTT GTCAAAGAGG ATGGACTAAA TGATCTCTGA				1259

AAGTGTTGAA G

(2) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 317 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: polypeptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

Met Ala Val Gln Gly Ser Gln Arg Arg Leu Leu Gly Ser Leu Asn Ser
 1           5           10           15
Thr Pro Thr Ala Ile Pro Gln Leu Gly Leu Ala Ala Asn Gln Thr Gly
          20           25           30
Ala Arg Cys Leu Glu Val Ser Ile Ser Asp Gly Leu Phe Leu Ser Leu
          35           40           45
Gly Leu Val Ser Leu Val Glu Asn Ala Leu Val Val Ala Thr Ile Ala
          50           55           60
Lys Asn Arg Asn Leu His Ser Pro Met Tyr Cys Phe Ile Cys Cys Leu
          65           70           75           80
Ala Leu Ser Asp Leu Leu Val Ser Gly Ser Asn Val Leu Glu Thr Ala
          85           90           95
Val Ile Leu Leu Leu Glu Ala Gly Ala Leu Val Ala Arg Ala Ala Val
          100          105          110
Leu Gln Gln Leu Asp Asn Val Ile Asp Val Ile Thr Cys Ser Ser Met
          115          120          125
Leu Ser Ser Leu Cys Phe Leu Gly Ala Ile Ala Val Asp Arg Tyr Ile
          130          135          140
Ser Ile Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Leu Pro Arg
          145          150          155          160
Ala Arg Arg Arg Val Ala Ala Ile Trp Val Ala Ser Val Val Phe Ser
          165          170          175
Thr Leu Phe Ile Ala Tyr Tyr Asp His Val Ala Val Leu Leu Cys Leu
          180          185          190
Val Val Phe Phe Leu Ala Met Leu Val Leu Met Ala Val Leu Tyr Val
          195          200          205
His Met Leu Ala Arg Ala Cys Gln His Ala Gln Gly Ile Ala Arg Leu
          210          215          220
His Lys Arg Gln Arg Pro Val His Gln Gly Phe Gly Leu Lys Gly Ala
          225          230          235          240

```

Val	Thr	Leu	Thr	Ile	Leu	Leu	Gly	Ile	Phe	Phe	Leu	Cys	Trp	Gly	Pro
				245					250					255	
Phe	Phe	Leu	His	Leu	Thr	Leu	Ile	Val	Leu	Cys	Pro	Glu	His	Pro	Thr
			260					265					270		
Cys	Gly	Cys	Ile	Phe	Lys	Asn	Phe	Asn	Leu	Phe	Leu	Ala	Leu	Ile	Ile
		275					280					285			
Cys	Asn	Ala	Ile	Ile	Asp	Pro	Leu	Ile	Tyr	Ala	Phe	His	Ser	Gln	Glu
	290					295					300				
Leu	Arg	Arg	Thr	Leu	Lys	Glu	Val	Leu	Thr	Cys	Ser	Trp			
305					310					315					

(2) INFORMATION FOR SEQ ID NO: 3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GGGAATTCTG TGTGTNATCN CNGTGGACCG GTA

33

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GGGGATCCGA AGAAGGGNAA CCAGCAGAGN ATGAA

35

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 285 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (PCR-fragment)

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..285

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

TTC	TAC	GCA	CTG	CGC	TAC	CAC	AGC	ATC	GTG	ACC	ATG	CGC	CGC	ACT	GTG	48
Phe	Tyr	Ala	Leu	Arg	Tyr	His	Ser	Ile	Val	Thr	Met	Arg	Arg	Thr	Val	
1				5				10						15		
GTG	GTG	CTT	ACG	GTC	ATC	TGG	ACG	TTC	TGC	ACG	GGG	ACT	GGC	ATC	ACC	96
Val	Val	Leu	Thr	Val	Ile	Trp	Thr	Phe	Cys	Thr	Gly	Thr	Gly	Ile	Thr	
		20						25					30			
ATG	GTG	ATC	TTC	TCC	CAT	CAT	GTG	CCC	ACA	GTG	ATC	ACC	TTC	ACG	TCG	144
Met	Val	Ile	Phe	Ser	His	His	Val	Pro	Thr	Val	Ile	Thr	Phe	Thr	Ser	
		35					40					45				
CTG	TTC	CCG	CTG	ATG	CTG	GTC	TTC	ATC	CTG	TGC	CTC	TAT	GTG	CAC	ATG	192
Leu	Phe	Pro	Leu	Met	Leu	Val	Phe	Ile	Leu	Cys	Leu	Tyr	Val	His	Met	
	50					55					60					
TTC	CTG	CTG	GCT	CGA	TCC	CAC	ACC	AGG	AAG	ATC	TCC	ACC	CTC	CCC	AGA	240
Phe	Leu	Leu	Ala	Arg	Ser	His	Thr	Arg	Lys	Ile	Ser	Thr	Leu	Pro	Arg	
65				70				75							80	
GCC	AAC	ATG	AAA	GGG	GCC	ATC	ACC	CTC	ACC	ATC	CTG	CTG	GGC	ATT		285
Ala	Asn	Met	Lys	Gly	Ala	Ile	Thr	Leu	Thr	Ile	Leu	Leu	Gly	Ile		
			85					90					95			

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 95 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: polypeptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Phe	Tyr	Ala	Leu	Arg	Tyr	His	Ser	Ile	Val	Thr	Met	Arg	Arg	Thr	Val	
1				5				10						15		
Val	Val	Leu	Thr	Val	Ile	Trp	Thr	Phe	Cys	Thr	Gly	Thr	Gly	Ile	Thr	
		20						25					30			
Met	Val	Ile	Phe	Ser	His	His	Val	Pro	Thr	Val	Ile	Thr	Phe	Thr	Ser	
		35					40					45				
Leu	Phe	Pro	Leu	Met	Leu	Val	Phe	Ile	Leu	Cys	Leu	Tyr	Val	His	Met	
	50					55				60						
Phe	Leu	Leu	Ala	Arg	Ser	His	Thr	Arg	Lys	Ile	Ser	Thr	Leu	Pro	Arg	
65				70				75							80	
Ala	Asn	Met	Lys	Gly	Ala	Ile	Thr	Leu	Thr	Ile	Leu	Leu	Gly	Ile		
			85					90					95			

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 306 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (PCR-fragment)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..306

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

TTC TAC GCA CTG CGC TAC CAC AGC ATC GTG ACG GCG AGG CGC TCA GGG	48
Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Ala Arg Arg Ser Gly	
1 5 10 15	
GCC ATC ATC GCC GGC ATC TGG GCT TTC TGC ACG GGC TGC GGC ATT GTC	96
Ala Ile Ile Ala Gly Ile Trp Ala Phe Cys Thr Gly Cys Gly Ile Val	
20 25 30	
TTC ATC CTG TAC TCA GAA TCC ACC TAC GTC ATC CTG TGC CTC ATC TCC	144
Phe Ile Leu Tyr Ser Glu Ser Thr Tyr Val Ile Leu Cys Leu Ile Ser	
35 40 45	
ATG TTC TTC GCT ATG CTG TTC CTC CTG GTG TCT CTG TAC ATA CAC ATG	192
Met Phe Phe Ala Met Leu Phe Leu Leu Val Ser Leu Tyr Ile His Met	
50 55 60	
TTC CTC CTG GCG CGG ACT CAC GTC AAG CGG ATC GCG CTC TGC CCG GGG	240
Phe Leu Leu Ala Arg Thr His Val Lys Arg Ile Ala Leu Cys Pro Gly	
65 70 75 80	
CCA GCT CTG CGC GGC AGA GGA CCA GCA TGC AGG GGC GCG GTC ACC CTC	288
Pro Ala Leu Arg Gly Arg Gly Pro Ala Cys Arg Gly Ala Val Thr Leu	
85 90 95	
ACC ATC CTG CTG GGC ATT	306
Thr Ile Leu Leu Gly Ile	
100	

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: polypeptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Ala Arg Arg Ser Gly
1 5 10 15

[illegible]

(2) INFORMATION FOF SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 312 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (PCR-fragment)

(ix) FEATURE:

- (A) NAME/KEY: CDS
(B) LOCATION: 1..312

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

TTC	TAC	GCA	CTG	CGT	TAC	CAC	AGC	ATC	GTG	ACC	GTG	CGG	CGG	GCC	CTC	48
Phe	Tyr	Ala	Leu	Arg	Tyr	His	Ser	Ile	Val	Thr	Val	Arg	Arg	Ala	Leu	
1				5					10					15		
ACC	TTG	ATC	GTG	GCC	ATC	TGG	GTC	TGC	TGC	GGC	GTC	TGT	GGC	GTG	GTG	96
Thr	Leu	Ile	Val	Ala	Ile	Trp	Val	Cys	Cys	Gly	Val	Cys	Gly	Val	Val	
			20					25					30			
TTC	ATC	GTC	TAC	TCG	GAG	AGC	AAA	ATG	GTC	ATT	GTG	TGC	CTC	ATC	ACC	144
Phe	Ile	Val	Tyr	Ser	Glu	Ser	Lys	Met	Val	Ile	Val	Cys	Leu	Ile	Thr	
		35					40					45				
ATG	TTC	TTC	GCC	ATG	ATG	CTC	CTC	ATG	GGC	ACC	CTC	TAC	GTG	CAC	ATG	192
Met	Phe	Phe	Ala	Met	Met	Leu	Leu	Met	Gly	Thr	Leu	Tyr	Val	His	Met	
	50					55					60					
TTC	CTC	TTT	GCG	CGG	CTG	CAC	GTC	AAG	CGC	ATA	GCA	GCA	CTG	CCA	CCT	240
Phe	Leu	Phe	Ala	Arg	Leu	His	Val	Lys	Arg	Ile	Ala	Ala	Leu	Pro	Pro	
65					70				75					80		
GCC	GAC	GGG	GTG	GCG	CCA	CAG	CAA	CAC	TCA	TGC	ATG	AAG	GGG	GCA	GTC	288
Ala	Asp	Gly	Val	Ala	Pro	Gln	Gln	His	Ser	Cys	Met	Lys	Gly	Ala	Val	
				85					90					95		

ACC CTC ACC ATC CTG CTG GGC ATT
 Thr Leu Thr Ile Leu Leu Gly Ile
 100

312

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 104 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: polypeptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Phe Tyr Ala Leu Arg Tyr His Ser Ile Val Thr Val Arg Arg Ala Leu
 1 5 10 15
 Thr Leu Ile Val Ala Ile Trp Val Cys Cys Gly Val Cys Gly Val Val
 20 25 30
 Phe Ile Val Tyr Ser Glu Ser Lys Met Val Ile Val Cys Leu Ile Thr
 35 40 45
 Met Phe Phe Ala Met Met Leu Leu Met Gly Thr Leu Tyr Val His Met
 50 55 60
 Phe Leu Phe Ala Arg Leu His Val Lys Arg Ile Ala Ala Leu Pro Pro
 65 70 75 80
 Ala Asp Gly Val Ala Pro Gln Gln His Ser Cys Met Lys Gly Ala Val
 85 90 95
 Thr Leu Thr Ile Leu Leu Gly Ile
 100

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 372 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (PCR-fragment)

(ix) FEATURE:

- (A) NAME/KEY: CDS
 (B) LOCATION: 1..372

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

CTG TGT GTG ATC GCG CTG GAC CGG TAC ATC TCC ATC TTC TAC GCA CTG
 Leu Cys Val Ile Ala Leu Asp Arg Tyr Ile Ser Ile Phe Tyr Ala Leu
 1 5 10 15

48

CGC TAC CAC AGC ATC GTG ACC CTG CCG CGG GCG CCG GAA GCC GTT GCG	96
Arg Tyr His Ser Ile Val Thr Leu Pro Arg Ala Pro Glu Ala Val Ala	
20 25 30	
GCC ATC TGG GTG GCC AGT GTC GTC TTC AGC ACG CTC TTC ATC GCC TAC	144
Ala Ile Trp Val Ala Ser Val Val Phe Ser Thr Leu Phe Ile Ala Tyr	
35 40 45	
TAC GAC CAC GTG GCC GTC CTG CTG TGC CTC GTG GTC TTC TTC CTG GCT	192
Tyr Asp His Val Ala Val Leu Leu Cys Leu Val Val Phe Phe Leu Ala	
50 55 60	
ATG CTG GTG CTC ATG GCC GTG CTG TAC GTC CAC ATG CTG GCC CGG GCC	240
Met Leu Val Leu Met Ala Val Leu Tyr Val His Met Leu Ala Arg Ala	
65 70 75 80	
TGC CAG CAC GCC CAG GGC ATC GCC CGG CTC CAC AAG AGG CAG CGC CCG	288
Cys Gln His Ala Gln Gly Ile Ala Arg Leu His Lys Arg Gln Arg Pro	
85 90 95	
GTC CAC CAG GGC TTT GGC CTT AAA GGC GCT GTC ACC CTC ACC ATC CTG	336
Val His Gln Gly Phe Gly Leu Lys Gly Ala Val Thr Leu Thr Ile Leu	
100 105 110	
CTG GGC ATT TTC ACC GTC TCG TGG CGC CCC TTC TTC	372
Leu Gly Ile Phe Thr Val Ser Trp Arg Pro Phe Phe	
115 120	

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Leu Cys Val Ile Ala Leu Asp Arg Tyr Ile Ser Ile Phe Tyr Ala Leu	
1 5 10 15	
Arg Tyr His Ser Ile Val Thr Leu Pro Arg Ala Pro Glu Ala Val Ala	
20 25 30	
Ala Ile Trp Val Ala Ser Val Val Phe Ser Thr Leu Phe Ile Ala Tyr	
35 40 45	
Tyr Asp His Val Ala Val Leu Leu Cys Leu Val Val Phe Phe Leu Ala	
50 55 60	
Met Leu Val Leu Met Ala Val Leu Tyr Val His Met Leu Ala Arg Ala	
65 70 75 80	
Cys Gln His Ala Gln Gly Ile Ala Arg Leu His Lys Arg Gln Arg Pro	
85 90 95	

Leu Gly Ile Phe Thr Val Ser Trp Arg Pro Phe Phe
115 120

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

35

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 34 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

34

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1650 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

```
(ix) FEATURE:
      (A) NAME/KEY: CDS
      (B) LOCATION: 616..1590
```

CTTGAGGAGA	ATGTCGTGCA	GTAGCCTTAG	GAATGTGAAC	ATTGGGAGAC	TGGCTGGGAT	60
TTTGTAGGTT	ATGAGAAAGG	GACACTTATG	ATATGTGAAC	TTGAGCCCG	GAGAGAAGCC	120
ATAAAAAGTG	AAACTGTCCT	GGGCACTTGG	AGGTGAGTGT	CTCTCTAGTA	AGATGCATGT	180

GAAAGGCCTG GGAGCTGAAA GCAAGGAGAG CAGAAGAGGC TGGTGAAGAT TCTAATCTGC	240
GTGTCCAGGG GCACTCTTCC AGGTCTCAGG AACGCAGGTC AGAATGTGCA AGCCAGCTGC	300
CGGGCACGTG GCTCACCCT GTAGTACCAG CACTTTGGGA GGCTGAGAGA GAAGATCGCT	360
TGTGGCCAGG AGTTTGAGAC CAGACTGGGG CTTCATAGGG AGACCCTGTC TCTTAAAAAA	420
AAAAAAAAAA AAGGACTGAG TGAGCCGAGC CCAGTCCTCT CATGCACTGT GTCATTTCATC	480
CCCTTTCTTA GGCTGTGTTG GTTCTAGGCT AGCTGCTGTC TTTCTTTGGT AGGCTGCTAA	540
CCTCTTTGGA TTGTGAATTT AAAACATGTT TTACAGTAAA TTTGCTGCCA AGACAAGAGG	600
TGTATTTCTC CAGCA ATG AAT TCC TCA TTT CAC CTG CAT TTC TTG GAT CTC	651
Met Asn Ser Ser Phe His Leu His Phe Leu Asp Leu	
1 5 10	
AAC CTG AAT GCC ACA GAG GGC AAC CTT TCA GGA CCC AAT GTC AAA AAC	699
Asn Leu Asn Ala Thr Glu Gly Asn Leu Ser Gly Pro Asn Val Lys Asn	
15 20 25	
AAG TCT TCA CCA TGT GAA GAC ATG GGC ATT GCT GTG GAG GTG TTT CTC	747
Lys Ser Ser Pro Cys Glu Asp Met Gly Ile Ala Val Glu Val Phe Leu	
30 35 40	
ACT CTG GGT GTC ATC AGC CTC TTG GAG AAC ATC TTG GTC ATA GGG GCC	795
Thr Leu Gly Val Ile Ser Leu Leu Glu Asn Ile Leu Val Ile Gly Ala	
45 50 55 60	
ATA GTG AAG AAC AAA AAC CTG CAC TCC CCC ATG TAC TTC TTC GTG TGC	843
Ile Val Lys Asn Lys Asn Leu His Ser Pro Met Tyr Phe Phe Val Cys	
65 70 75	
AGC CTG GCA GTG GCG GAC ATG CTG GTG AGC ATG TCC AGT GCC TGG GAG	891
Ser Leu Ala Val Ala Asp Met Leu Val Ser Met Ser Ser Ala Trp Glu	
80 85 90	
ACC ATC ACC ATC TAC CTA CTC AAC AAC AAG CAC CTA GTG ATA GCA GAC	939
Thr Ile Thr Ile Tyr Leu Leu Asn Asn Lys His Leu Val Ile Ala Asp	
95 100 105	
GCC TTT GTG CGC CAC ATT GAC AAT GTG TTT GAC TCC ATG ATC TGC ATT	987
Ala Phe Val Arg His Ile Asp Asn Val Phe Asp Ser Met Ile Cys Ile	
110 115 120	
TCC GTG GTG GCA TCC ATG TGC AGC TTA CTG GCC ATT GCA GTG GAT AGG	1035
Ser Val Val Ala Ser Met Cys Ser Leu Leu Ala Ile Ala Val Asp Arg	
125 130 135 140	
TAC GTC ACC ATC TTC TAC GGC CTG CGC TAC CAC CAC ATC ATG ACG GCG	1083
Tyr Val Thr Ile Phe Tyr Ala Leu Arg Tyr His His Ile Met Thr Ala	
145 150 155	
AGG CGC TCA GGG GCC ATC ATC GCC GGC ATC TGG GCT TTC TGC ACG GGC	1131
Arg Arg Ser Gly Ala Ile Ile Ala Gly Ile Trp Ala Phe Cys Thr Gly	
160 165 170	

TGC GGC ATT GTC TTC ATC CTG TAC TCA GAA TCC ACC TAC GTC ATC CTG	1179
Cys Gly Ile Val Phe Ile Leu Tyr Ser Glu Ser Thr Tyr Val Ile Leu	
175 180 185	
TGC CTC ATC TCC ATG TTC TTC GCT ATG CTG TTC CTC CTG GTG TCT CTG	1227
Cys Leu Ile Ser Met Phe Phe Ala Met Leu Phe Leu Leu Val Ser Leu	
190 195 200	
TAC ATA CAC ATG TTC CTC CTG GCG CGG ACT CAC GTC AAG CGG ATC GCG	1275
Tyr Ile His Met Phe Leu Leu Ala Arg Thr His Val Lys Arg Ile Ala	
205 210 215 220	
CTC TGC CCG GGG CCA GCT CTG CGC GGC AGA GGA CCA GCA TGG CAG GGC	1323
Leu Cys Pro Gly Pro Ala Leu Arg Gly Arg Gly Pro Ala Trp Gln Gly	
225 230 235	
GCG GTC ACC GTC ACC ATG CTG CTG GGC GTG TTT ACC GTG TGC TGG GCG	1371
Ala Val Thr Val Thr Met Leu Leu Gly Val Phe Thr Val Cys Trp Ala	
240 245 250	
CCG TTC TTC CTT CAT CTC ACT TTA ATG CTT TCT TGC CCT CAG AAC CTC	1419
Pro Phe Phe Leu His Leu Thr Leu Met Leu Ser Cys Pro Gln Asn Leu	
255 260 265	
TAC TGC TCT CGC TTC ATG TCT CAC TTC AAT ATG TAC CTC ATA CTC ATC	1467
Tyr Cys Ser Arg Phe Met Ser His Phe Asn Met Tyr Leu Ile Leu Ile	
270 275 280	
ATG TGT AAT TCC GTG ATG GAC CCT CTC ATA TAT GCC TTC CGC AGC CAA	1515
Met Cys Asn Ser Val Met Asp Pro Leu Ile Tyr Ala Phe Arg Ser Gln	
285 290 295 300	
GAG ATG CGG AAG ACC TTT AAG GAG ATT ATT TGC TGC CGT GGT TTC AGG	1563
Glu Met Arg Lys Thr Phe Lys Glu Ile Ile Cys Cys Arg Gly Phe Arg	
305 310 315	
ATC GCC TGC AGC TTT CCC AGA AGG GAT TAACGACAAA GTGCTCCTCT	1610
Ile Ala Cys Ser Phe Pro Arg Arg Asp	
320 325	
CTGTGGCTCT GTTCTCCTTT GTTTGCTCAC CTATGACAAA	1650

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 325 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Met	Asn	Ser	Ser	Phe	His	Leu	His	Phe	Leu	Asp	Leu	Asn	Leu	Asn	Ala
1					5				10					15	
Thr	Glu	Gly	Asn	Leu	Ser	Gly	Pro	Asn	Val	Lys	Asn	Lys	Ser	Ser	Pro

20					25					30					
Cys	Glu	Asp	Met	Gly	Ile	Ala	Val	Glu	Val	Phe	Leu	Thr	Leu	Gly	Val
		35					40					45			
Ile	Ser	Leu	Leu	Glu	Asn	Ile	Leu	Val	Ile	Gly	Ala	Ile	Val	Lys	Asn
	50					55					60				
Lys	Asn	Leu	His	Ser	Pro	Met	Tyr	Phe	Phe	Val	Cys	Ser	Leu	Ala	Val
	65					70					75				80
Ala	Asp	Met	Leu	Val	Ser	Met	Ser	Ser	Ala	Trp	Glu	Thr	Ile	Thr	Ile
				85					90					95	
Tyr	Leu	Leu	Asn	Asn	Lys	His	Leu	Val	Ile	Ala	Asp	Ala	Phe	Val	Arg
			100					105					110		
His	Ile	Asp	Asn	Val	Phe	Asp	Ser	Met	Ile	Cys	Ile	Ser	Val	Val	Ala
		115					120					125			
Ser	Met	Cys	Ser	Leu	Leu	Ala	Ile	Ala	Val	Asp	Arg	Tyr	Val	Thr	Ile
	130					135					140				
Phe	Tyr	Ala	Leu	Arg	Tyr	His	His	Ile	Met	Thr	Ala	Arg	Arg	Ser	Gly
	145					150					155				160
Ala	Ile	Ile	Ala	Gly	Ile	Trp	Ala	Phe	Cys	Thr	Gly	Cys	Gly	Ile	Val
				165					170					175	
Phe	Ile	Leu	Tyr	Ser	Glu	Ser	Thr	Tyr	Val	Ile	Leu	Cys	Leu	Ile	Ser
			180					185					190		
Met	Phe	Phe	Ala	Met	Leu	Phe	Leu	Leu	Val	Ser	Leu	Tyr	Ile	His	Met
		195					200					205			
Phe	Leu	Leu	Ala	Arg	Thr	His	Val	Lys	Arg	Ile	Ala	Leu	Cys	Pro	Gly
	210					215					220				
Pro	Ala	Leu	Arg	Gly	Arg	Gly	Pro	Ala	Trp	Gln	Gly	Ala	Val	Thr	Val
	225					230					235				240
Thr	Met	Leu	Leu	Gly	Val	Phe	Thr	Val	Cys	Trp	Ala	Pro	Phe	Phe	Leu
				245					250					255	
His	Leu	Thr	Leu	Met	Leu	Ser	Cys	Pro	Gln	Asn	Leu	Tyr	Cys	Ser	Arg
			260					265					270		
Phe	Met	Ser	His	Phe	Asn	Met	Tyr	Leu	Ile	Leu	Ile	Met	Cys	Asn	Ser
			275				280					285			
Val	Met	Asp	Pro	Leu	Ile	Tyr	Ala	Phe	Arg	Ser	Gln	Glu	Met	Arg	Lys
	290					295					300				
Thr	Phe	Lys	Glu	Ile	Ile	Cys	Cys	Arg	Gly	Phe	Arg	Ile	Ala	Cys	Ser
	305					310					315				320
Phe	Pro	Arg	Arg	Asp											
				325											

(2) INFORMATION FOR SEQ ID NO: 17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO 17:

GGAAGCTTTC TTTGGTAGGC TG

22

(2) INFORMATION FOR SEQ ID NO: 18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGTCTAGAGC CACAGAGAGG AG

22

(2) INFORMATION FOR SEQ ID NO: 19:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CTGCATTTCT TGGATCT

17

(2) INFORMATION FOR SEQ ID NO: 20:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 17 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (synthetic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

AAGCTGCACA TGGATGC

17